METHOD FOR IMPLEMENTING CONTENT DELIVERY NETWORK (CDN) INT...

CROSS REFERENCE TO RELATED APPLICATIONS

This application is the US national phase of PCT application PCT/EP2003/04059, filed 17 April 2003, published 30 October 2007 as WO 2003/090423, and claiming the priority of Italian patent application TO2002A000341 itself filed 19 April 2002, whose entire disclosures are herewith incorporated by reference.

FIELD OF THE INVENTION

This invention relates in general to techniques generally known as "internetworking".

BACKGROUND OF THE INVENTION

In general terms, the basic objective of internetworking is the co-operation and interoperability of elementary systems (seen as "black boxes") to create a macrosystem capable of presenting the characteristics of the constituent systems with the addition of a number of advantages.

First, when two or more different administrative entities reach an internetworking agreement they extend (within contractual limits) their respective catchment areas without additional expenses for wiring or structural purposes. It is reasonable to think that the service quality perceived by the final user may be increased due to the larger size of the reference network.

In the specific case of the so-called Content Delivery Networks, or CDNs, additional contents and diversification is also provided.

For its very nature, an internetworking solution requires the presence of interface components which, on elementary system (i.e. single CDN) side have a complete overview of the evolution of the static and dynamic characteristics and which on the "rest of the world" side (i.e. on the side of the other internetworking networks, that is on peer side) have only the comprehensive information needed to establish profitable intersystem communication. The term "profitable" here means efficient, safe and reliable being provided with the mechanisms that this type of solution entails.

Regulations concerning the matter are currently being drawn up, as documented, for example, by the following draft standards published by IETF (Internet Engineering Task Force) which can be retrieved from the organization's web site, namely: "A Model for CDN Peering", by M. Day, B. Cain, G. Tomlinson and P. Rzewski, May 2001; "Content Internetworking Architectural Overview", by M. Green, B. Cain, G. Tomlinson, S. Thomas e P. Rzewski, March 2001; "Known Mechanisms for Content Internetworking", by F. Douglis, I. Chaudhri and P. Rzewski, November 2001.

The interface components are called Content Internetworking Gateways or CIGs. CIGs have a complete overview of the environment within their respective CDN and perceive the data related to remote environments through protocols for exchanging data, called "advertisement".

From an abstract point of view, a CIG must route requests (i.e. perform request-routing functions), on the basis of all data from the pre-existing infra-CDN modules (distribution system and monitoring system) and equivalent remote devices.

According to the aforesaid standards, the CIG routes a client's content requests.

Specifically, having received a request for a certain content and having verified that the content is available in its respective CDN, the CIG sends the corresponding required content cache ID to the client. The concerned CIG is consequently capable of routing the request, also when the content is hosted in the cache of another CDN.

In this situation, when several CDN networks are internetworking, the CIGs perform address resolution and content request-routing functions, which on internet level are remitted to other network members, particularly by involving the so-called DNS (Directory Name Service or Domain Name Server).

This leads to splitting/duplication of functions which causes several problems. The problems are related, among other, to the need of ensuring correct synchronisation between CIGs and devices in the Internet and to the fact that the request from a certain client is processed differently according to whether the request involves the CDN level or not.

OBJECT OF THE INVENTION

The object of the invention is to overcome these shortcomings.

SUMMARY OF THE INVENTION

The object is obtained according to the invention thanks to a procedure basically comprising the steps of collecting in the interface components CIG routing data related to the association of the contents and the caches which contain them; and transferring the routing data from at least one of the interface components to the Directory Name Service or Domain Name

Server of the respective network, so that access by the client of the respective network of contents of the networks in the set of CDN is implemented through the Directory Name Service or Domain Name Server (DNS) of the network. The invention also relates to a corresponding system of internetworking CDN networks and a respective interface or CIG component.

BRIEF DESCRIPTION OF DRAWINGS

The invention will now be described, by way of example only, with reference to the accompanying drawings wherein:

FIG. 1 generally illustrates the internetworking organization criteria of two CDN networks,

FIG. 2 generally illustrates the structures of a Content Internetworking Gateway, or CIG, according to the invention,

FIGS. 3 and 4 illustrate different infra-CDN and inter-CDN request-routing methods,

FIGS. 5 to 7 illustrate the typical CIG context diagrams at various levels detail according to the invention,

FIG. 8 shows the finite state automaton of a corresponding CIG,

FIG. 9 is a time diagram showing the opening of a corresponding session, and

FIGS. 10 to 13 illustrate the format of the various messages exchanged by a CIG according to the invention.

BEST MODE FOR CARRYING OUT THE INVENTION

The diagram in FIG. 1 illustrates the collocation of two Content Internetworking Gateways (hereinafter called CIG for short) intended to permit exchange of "advertizement" data in the

context of a set formed by two Content Delivery Networks CDN1 and CDN2 in combination with an Origin Server (OS) each.

Each CDN shown here consists of a respective administrative domain with a Directory Name Service or Domain Name Server (DNS for short), management center, cache memories and connections to client function terminals.

FIG. 1 shows the role of the CIGs in the internetworking process. One of the specific characteristics of a CIG is the degree (or level) of integration, as a parameter, respect to the modules which are already present and operational within a CDN.

The higher or lower efficiency of the respective interface functions can be assessed according to this parameter.

FIG. 2 briefly illustrates the interface components which form a CIG according to the invention in the currently preferred form of embodiment.

Specifically, the concerned CIG consists of:

a first interface module, called Request-Routing Interface or RRI, which exchanges data with the remote CIGs according to CNAP protocol specifications (described in detail in the description that follows),

a second interface module, called DNS Interface or DNSI, which interfaces with the DNS of the CDN to which the CIG belongs,

a third interface module, called Distribution Information Interface, or DII, which retrieves data on the availability of contents from the distribution system of the CDN to which the CIG belongs,

a forth interface module, called Monitoring Information Interface, or MII, which interacts with the monitoring system, and finally

a central module, called Request-Routing Process, or RRP, which collects and processes the information received to implement the request-routing function: the latter module is the CIG core.

It is noted that the aforesaid architecture, although preferred, is not absolutely imperative or binding, at least as concerns the presence of the third or fourth interface module described above.

Further reference to the CNAP protocol may be found in

 "Content Network Advertizement Protocol (CNAP) "by B. Cain, O. Spatscheck, L. Amini, A. Barbir, M. May and D. Kaplan, November 2001, which may also be retrieved from the IETF web site.

Briefly, the CIG consists of a central module which is the "brain" of the device and a certain number of interface modules between the CIG and the pre-existing infrastructure.

The described request-routing technique solution refers to modules implementing DNS technology.

Consequently, two likely internetworking scenarios may be hypothesized and illustrated in an event trace diagram.

FIG. 3 shows a classical content routing scenario, so to speak, the term herein indicating a standard routing process in a CDN (implementing DNS technology) in which DNS table updating is delegated to the CIG by means of the DNSI module.

Extending the example to an actual internetworking case is easy with this procedure.

The labels and the directions of the arrows in the figures help to understand the real sequence of events: a user makes a content request to the DNS system which works in standard mode. The DNS responds with the best surrogate IP address according to the routing policies applied at the time. The CIG periodically updates the DNS tables, according to the information received from the distribution and monitoring system; note that in this first case, the system is "isolated", so to speak.

Conversely, in the situation in FIG. 4, the selected surrogate servers belong to another administrative domain, i.e. CDN. The dynamics appears essentially similar to the example above. In this case, as above, the client queries the DNS which replies with the best surrogate server IP address. The difference is in the data retrieving and updating method. The bi-directional arrows in the central section indicate the periodical exchange of routing data between entities on the same hierarchic level (peers), i.e. the CIGs, according to the conventions and the specifications of the CNAP protocol. This type of data is similar to infra-CDN data, and used by a CIG to update the DNS tables on the basis of a wider range of data with respect to that which occurs in known architectures.

The roles of the modules which form a CIG operating according to the invention will now be described with reference to the diagram indicated by the acronym DFD (Data Flow Diagram).

The higher level approach consists in the use of a so-called context diagram. The diagram represents the interactions between the whole CIG and the "outside world".

As shown, the CIG appears as a single entity capable of interacting with the rest of the world.

The CIG routes requests according to the information from the other entities with which it communicates. More in detail, it receives data from peers, from the distribution system and the local monitoring system. After processing, the data, the DNS tables and the log file archives are updated.

At this point, the request-routing system can be observed closer, by splitting into subsystems and representing the functions on different levels of detail.

Two subsequent expansions are illustrated in FIG. 6 and FIG. 7.

Specifically, the interface processes clearly appear in FIG. 6, corresponding to a first level of detail: these are "buffer" modules which communicate with the central process on one side and with the outside world on the other.

FIG. 7, on the other hand, illustrates the functions of the RRP core. The RRP receives data from the rest of the world and transmits them via the interfaces, extracts useful information on cache and/or content state, evaluates the need to update and consequently modifies its own database, the DNS tables and the log file archives.

Finally, if required, it sends the message to its peers, through the request-routing interface RRI.

The request-routing interface RRI interfaces with the rest of the world. From this point of view, it is the most important module in the internetworking procedure, because it is directly implied in inter-CDN communication; as mentioned above, this communication is carried out according to the conventions of the CNAP protocol which was designed for this purpose.

This module is responsible for translating the messages from CNAP (inter-CDN) format to a format which can be understood by the CIG (infra-CDN) central process, or RRP. The CNAP protocol requires initial specifications (and periodical updating) or a set of data, which are static so to speak, referred to the internetworking system topology characteristics. For example, the following may be requested:

the local CNAS ID (i.e. the CDN to which the concerned CIG belongs);

the IP address of the local CIG computer;

the CNAS INFORMATION DISCLOSURE STATEMENT of remote interconnected systems (peers) with which internetworking will be established;

the IP addresses of the remote CIG computers corresponding to the systems mentioned in the point above;

the inter-CNAS level of confidences; and - a numeric coefficient indicating the "weight" in static conditions of each connection (practically similar to the geographical distance of the connection).

The protocol offers the possibility of diversifying the contractual internetworking relations with the introduction of level of confidences. In other words, before disseminating information on availability of a content to a remote CIG, the local CIG verifies whether the CIG is enabled to received the information according to the stipulated contact.

The CNAP connections, as required by the IETF for internetworking protocols, implement a reliable connection-oriented protocol on transportation level: for example, TCP

(Transmission Control Protocol), currently employed in Internet contexts, may be used.

The logical operations needed to establish a CNAP session are shown below.

This is carried out with specific reference to the finite state automaton diagram of the CIG as illustrated in FIG. 8.

During the initial state of the CIG, called IDLE, there is no CNAP session and no entity has intervened to change this situation. When the CIG intends to establish a CNAP session with a remote CIG, it sends an OPEN message and goes to OPENSENT state.

The remote, also initially in IDLE state, receives a request to open a CNAP session. It replies with a KEEPALIVE message and goes to OPENCONFIRM state.

Two cases may occur :

In the first case, the original CIG receives the KEEPALIVE message within a predetermined lapse of time: in this case, it goes to READY state and waits to send advertizement messages (i.e. messages carrying useful data, not only metadata, as initialization messages).

In the second case, the predetermined time-out expires before the original CIG receives the expected KEEPALIVE message: in this case, it returns to IDLE state and the communication attempt fails. In general, a NOTIFY message will inform the parties of the anomaly.

The remote CIG, having sent the following KEEPALIVE message, also goes to READY state and listens out for advertisement messages to be received.

The reception of a NOTIFY message makes the CIG go to IDLE state. As may easily be assumed from the description above, the CNAP connection is active and efficient if both involved CIG are in READY state.

FIG. 9 shows the sequences of state which characterize opening of a CNAP session and highlight the evolution of events in time.

The messages exchanged by the RRP core and the request-routing interface RRI may have the format shown in FIG. 10.

The meaning of the message fields is shown below:

URL: is the URL identifying the content of the message;

IP: is the IP address of the cache which distributes the contents;

CNAS ID: is the ID of the CDN to which the cache belongs;

CACHE STATE: is the state of the cache;

CONTENT STATE: is the state of the content in the cache; TTL: is the life time of the routing data.

The monitoring interface MII is the module which implements the interface between the RRP core and the local monitoring system, i.e. referred to the CDN to which the CIG belongs. The data which must be transferred to the RRP refer to the CDN cache state; the term "state" here indicates quantification of the available resources.

In this perspective, the format of a message from the monitoring interface MII may assume the appearance shown in FIG. 11.

The message has five fields whose meaning is illustrated below:

IP: the IP address of the cache to which the message refers;

CPU: percentage of CPU used by the cache;

MEM: percentage of RAM used by the cache;

DISC: percentage of disc used by the cache;

USERS: percentage of the number of connected users (in relation to the maximum service capacity of the concerned cache).

The parameters are classical performance indicators which are used to assess the conditions of use of the cache. Messages of this type are passed to the RRP at regular intervals of time.

The DII distribution interface is the interface module between the distribution system and the RRP core of the CIG. The DII interface collects information on the presence and availability of the cache contents. FIG. 12 shows the format of a possible message of this type.

The meaning of the fields is shown below:

URL: is the URL identifying the content to which the message refers;

CACHE: is the list of IP addresses of the caches in which the content is available;

LEVEL OF CONFIDENCE: is the level of confidence of the content;

CONTENT AVAILABILITY: indicates whether the content is available or not;

CACHE STATE: is the status of the cache;

TTL: indicates the life time of the routing data.

Three levels of confidence can be associated to the contents, i.e.:

low-contents can be exchanged with all interconnected CDNs;

medium-contents can be exchanged only with CDNs which have subscribed a MEDIUM level confidence agreement with the CDN that owns the content; and

high-contents can be exchanged only with CDNs which have subscribed a HIGH level confidence agreement with the CDN that owns the content.

The DNS interface, indicated by DNSI, is the interface module which must communicate with the DNS server, to update the tables. A possible format of the message useful for this purpose is shown in FIG. 13.

The meaning of the respective fields is:

OP: indicates the operation to be carried out on the DNS table (two operations are available, "add" and "delete");

REG TYPE: indicates the type of register;

DOMAIN NAME: indicates the name of the domain to which the message refers;

IP: is the address of the best cache IP address to serve the domain above;

TTL: is the life time of the register.

Alternatively, the DOMAIN NAME field may contain the entire URL of the content to which the message refers. In this way, the DNS can directly identify the best cache for content delivery.

The request-routing module RRP is, as mentioned above, the core of the system. It is responsible for processing the data received from the aforesaid interface modules, updating the DNS tables if required via the DNSI interface and forwarding the data to the other CIGs through the RRI interface.

It is also responsible for managing the log file archive.

Given the need to enable the respective DNS to perform the address resolution function (to make content delivery factually "transparent" with respect to the presence of a set of internetworking CDN networks), the RRP core must have a data structure which will store the states of the local CDN and the remote CDNs. The data structure must collect and organize the data referred to contents available in the internetworking system context and to the caches capable of providing the contents. Data structure definition is periodically updated by the RRP module, in a different way according to the type of message which prompted the updating process on a case-by-case basis.

Naturally, numerous changes can be implemented to the construction and embodiments of the invention herein envisaged without departing from the scope of the present invention, as defined by the following claims.